

Manned

NSC-254



UNPUBLISHED PRELIMINARY DATA

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

~~NSC-254~~

RE-3

STATISTICAL FILTERING OF SPACE
NAVIGATION MEASUREMENTS

by
James E. Potter and Robert G. Stern

August 1963

FACILITY FORM 102

NSC-15413

(ACCESSION NUMBER) NSC-15413 (THRU)

(PAGES) 19 (COPIES) 2

(NASA CR OR TMX OR AD NUMBER) CR 69237 (CATEGORY) 2

GPO FRI. \$ _____

CFSTI PRICE (D) \$ _____

Hard copy (HC) 1.00

Microfiche (MF) 50

653 July 65

EXPERIMENTAL ASTRONOMY LABORATORY
 MASSACHUSETTS INSTITUTE OF TECHNOLOGY
 CAMBRIDGE 39, MASSACHUSETTS

**STATISTICAL FILTERING OF SPACE
NAVIGATION MEASUREMENTS**

August 1963

**Note: This report has also been issued
as AIAA Preprint No. 63-333, presented
at the AIAA Guidance and Control Con-
ference at Massachusetts Institute of
Technology August 12-14, 1963.**

**EXPERIMENTAL ASTRONOMY LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE 39, MASSACHUSETTS**

Prepared for Publication by Jackson & Moreland, Inc.

TABLE OF CONTENTS

<u>Section</u>		<u>Page</u>
1	Introduction	3
2	The State Vector	3
3	Methods of Designing Estimators	4
4	Optimization Criteria	4
5	Mathematical Model	4
6	Maximum Likelihood Theory	5
7	Gauss-Markov Theorem	6
8	Minimizing the Volume of the Error Ellipsoid	6
9	Minimizing Quadratic Forms in the Estimation Uncertainty	7
10	Biased and Unbiased Estimates	8

APPENDIX

A	10
B	10
C	11
D	11

LIST OF REFERENCES	12
------------------------------	----

STATISTICAL FILTERING OF SPACE NAVIGATION MEASUREMENTS¹

James E. Potter²

Robert G. Stern³

Massachusetts Institute of Technology

Abstract

The criteria that have been used for designing an estimator which computes the state vector of a space vehicle from redundant navigational measurements are reviewed and compared. It is shown that under the assumption of linearity all these criteria lead to the same estimator (or filter). If all measurement uncertainties have Gaussian distributions, it is shown that the estimate obtained from optimum filter theory is identical with the maximum likelihood estimate.

A proof is presented that the Bayes estimator, which is biased in favor of the initial conditions, leads to the same result that would be obtained from an unbiased estimator to which six pseudo-measurements, representing the initial conditions, have been added.

Although the theory has been applied in the specific area of space navigation, the results obtained are generally applicable to the estimation of multi-dimensional random variables.

Nomenclature

General Notation

Underlining a lower-case letter indicates that the letter represents a column vector.

An asterisk over a capital letter indicates that the letter represents a matrix.

Enclosing a matrix within two vertical bars indicates the determinant of the matrix.

Angular brackets $\langle \rangle$ indicate the average value of the bracketed quantity.

¹The analysis in this paper represents one phase of research carried out for the National Aeronautics and Space Administration under Contracts NAS 9-153 and NaG 254-62.

²Staff Engineer, Instrumentation Laboratory

³Staff Engineer, Experimental Astronomy Laboratory

Superscript T following a vector or matrix indicates the transpose of the vector or matrix.

Superscript -1 following a square matrix indicates the inverse of the matrix.

Symbols

$\overset{*}{A}$	symmetric matrix (in Section 7 and Appendix C)
$\overset{*}{A}$	r-by-s matrix (in Section 9)
$\overset{*}{A}$	positive semidefinite r-by-r matrix (in Appendices B and D)
$\overset{*}{A}_1$	r-by-r auxiliary matrix (in Appendix D)
\underline{b}	mean value of positive semidefinite quadratic form which is to be minimized
$\overset{*}{B}$	symmetric matrix (in Section 7 and Appendix C)
$\overset{*}{B}$	s-by-r matrix (in Section 9)
$\overset{*}{B}$	r-by-s matrix (in Appendix B)
$\overset{*}{B}$	positive semidefinite r-by-r matrix (in Appendix D)
$\overset{*}{C}_{ji}$	n-by-n state transition matrix
d_i	i-th diagonal element of $\overset{*}{D}$
$\overset{*}{D}$	diagonal matrix
$\overset{*}{D}_1$	auxiliary diagonal matrix
\underline{e}	error vector associated with the estimated state vector $\hat{\underline{x}}$
\underline{e}_B	error vector associated with $\hat{\underline{x}}_B$
$\overset{*}{E}$	covariance matrix associated with \underline{e}
$\overset{*}{E}_B$	covariance matrix associated with \underline{e}_B
$\overset{*}{E}_U$	covariance matrix associated with error in $\hat{\underline{x}}_U$
$\overset{*}{F}_A$	arbitrary unbiased filter
$\overset{*}{F}_B$	arbitrary biased filter
$\overset{*}{F}_B$	unbiased filter which has same covariance matrix as $\overset{*}{F}_B$

$\overset{*}{F}_{ML}$	maximum likelihood filter	p_k	vector equal to $\overset{*}{M} \underline{h}_k$ (in Appendix C)
$\overset{*}{F}_{ML}$	maximum likelihood filter which includes the fictitious measurements	q	number of dimensions of \underline{w}
$\overset{*}{F}_{OB}$	optimum biased filter	$\overset{*}{Q}$	positive semidefinite matrix
$\overset{*}{F}_R$	n-by-n residual matrix (in Section 10)	\underline{r}	position vector
$\overset{*}{F}_U$	arbitrary unbiased filter which includes the fictitious measurements	$\ddot{\underline{r}}$	inertial acceleration vector
\underline{g}	many-body gravitation vector	$\delta \underline{r}$	variation in position vector
$\overset{*}{G}$	3-by-3 matrix relating $\delta \ddot{\underline{r}}$ to $\delta \underline{r}$	$\delta \underline{r}_i$	$\delta \underline{r}$ evaluated at $t = t_i$
h	non-negative constant	$\delta \ddot{\underline{r}}$	variation in inertial acceleration vector
\underline{h}_k	unit vector in direction of k-th coordinate	$\overset{*}{R}$	q-by-n matrix relating \underline{w} to \underline{e}
$\overset{*}{H}$	k-by-n matrix relating \underline{m} to \underline{x}	\underline{s}_i	vector constituting i-th column of $\overset{*}{A}_1$ (in Appendix D)
$\overset{*}{H}_B$	k-by-n matrix relating \underline{m}_B to \underline{x}	t	time
$\overset{*}{H}_U$	(k + n)-by-n matrix relating \underline{m}_U to \underline{x}	tr	trace of a square matrix
$\overset{*}{I}$	identity matrix	\underline{u}	measurement uncertainty vector
$\overset{*}{J}$	n-by-k matrix equal to $\overset{*}{F}_A$ minus $\overset{*}{F}_{ML}$	\underline{u}_B	measurement uncertainty vector corresponding to \underline{m}_B
k	number of measurements being processed	\underline{u}_U	measurement uncertainty vector corresponding to \underline{m}_U
$L(\underline{x})$	likelihood function of \underline{x}	$\overset{*}{U}$	covariance matrix associated with \underline{u}
m_k	principal subdeterminant of matrix $\overset{*}{M}$	$\overset{*}{U}_B$	covariance matrix associated with \underline{u}_B
\underline{m}	true measurement deviation vector	$\overset{*}{U}_U$	covariance matrix associated with \underline{u}_U
$\tilde{\underline{m}}$	observed measurement deviation vector	$\delta \underline{v}_i$	velocity variation vector at $t = t_i$
\underline{m}_B	true measurement deviation vector for measurements actually made (in Section 10)	V	volume of error ellipsoid
$\tilde{\underline{m}}_B$	observed value of \underline{m}_B	\underline{w}	vector whose error ellipsoid is to be minimized
\underline{m}_U	measurement deviation vector which includes the fictitious measurements	$\overset{*}{W}$	covariance matrix associated with \underline{w}
$\tilde{\underline{m}}_U$	observed value of \underline{m}_U	x_i	i-th component of \underline{x}
M_{ij}	element in i-th row and j-th column of $\overset{*}{M}$	\underline{x}	true state vector
$\overset{*}{M}$	symmetric matrix (in Appendix A)	\underline{x}_i	true state vector evaluated at $t = t_i$
$\overset{*}{M}$	nonsingular matrix used to diagonalize $\overset{*}{A}$ and $\overset{*}{B}$ (in Appendix C)	$\hat{\underline{x}}$	estimated state vector
$\overset{*}{M}$	orthogonal matrix which diagonalizes $\overset{*}{A}$ (in Appendix D)	$\hat{\underline{x}}_B$	biased estimate of \underline{x}
n	number of dimensions of \underline{x}	$\hat{\underline{x}}_U$	unbiased estimate of \underline{x} which includes the fictitious measurements
$\overset{*}{N}$	inverse of $\overset{*}{M}$ (in Appendix C)	\underline{y}	r-dimensional vector equal to $\overset{*}{B}\underline{z}$
$\overset{*}{O}$	zero matrix	\underline{z}	arbitrary s-dimensional vector
$p(\underline{u})$	probability density of \underline{u}	$\underline{0}$	zero vector
$p(\tilde{\underline{m}} \underline{x})$	probability density of $\tilde{\underline{m}}$ given \underline{x}	$\Gamma ()$	gamma function of the argument
$p(\underline{u} \underline{x})$	probability density of \underline{u} given \underline{x}		

Subscripts

- A relating to filter \hat{F}_A^*
ML relating to maximum likelihood filter \hat{F}_{ML}^*
OB relating to optimum biased filter \hat{F}_{OB}^*

1. Introduction

The vector equation of motion of a space vehicle in free flight in a many-body gravitational field is

$$\ddot{\underline{r}} = \underline{g}(\underline{r}, t) \quad (1-1)$$

The inertial acceleration $\ddot{\underline{r}}$ is due entirely to the many-body gravitation vector \underline{g} , which is a function of both spatial position and time.

In the usual guidance analysis Equation (1-1) is linearized by assuming that the vehicle's actual position and velocity at any specified time differ only slightly from the pre-computed nominal, or reference, position and velocity at that time. If the symbol δ is used to signify the variation of a quantity from its nominal value, the linearized equations of motion of the vehicle can be written in matrix form as follows:

$$\delta \ddot{\underline{r}} = \hat{G} \delta \underline{r} \quad (1-2)$$

Here $\delta \ddot{\underline{r}}$ and $\delta \underline{r}$ are three-component column vectors and \hat{G} is a 3-by-3 matrix which is shown in Reference (1) to be symmetric for any gravitational field. The elements of \hat{G} are functions of the characteristics of the reference trajectory and are independent of the components of $\delta \underline{r}$.

Equation (1-2) represents three second-order linear differential equations with variable coefficients. The solution of the system contains six constants of integration. When the mathematical model is specified and the reference trajectory has been pre-computed, the motion of the vehicle along its actual trajectory can be completely determined once the six integration constants have been evaluated. The object of this paper is to examine techniques for estimating the integration constants from a set of redundant measurements, each of which is corrupted by noise.

2. The State Vector

The integration constants are conveniently grouped in the six-component column vector \underline{x}_i ,

known as the "state vector," which consists of the three components of $\delta \underline{r}_i$ and the three components of the velocity variation vector $\delta \underline{v}_i$ at some specified time t_i .

$$\underline{x}_i = \begin{bmatrix} \delta \underline{r}_i \\ \delta \underline{v}_i \end{bmatrix} \quad (2-1)$$

Various values of t_i have been used in the literature to define the particular state vector in terms of which the actual motion of the vehicle is expressed. For example, in Reference (2) t_i is the time of cut-off of the main propulsion system and thus the time of initiation of the free-fall trajectory; in References (3) and (4) t_i is the "current" time, that is, the time at which the most recent measurement was made or the time at which the next measurement is to be made; and in Reference (5) t_i is the nominal time of arrival at the destination.

The goal of the data processing is to minimize the errors in the estimate of the components of \underline{x}_i . It is shown in the later sections of this paper that, for the estimation techniques considered, the filter which minimizes the r. m. s. uncertainties in the components of the state vector at one specified time minimizes the r. m. s. uncertainties in the components of the state vector at any other time.

If position variation is more significant than velocity variation, as is the case for the destination state vector when the mission objective is to pass through a fixed point in inertial space at a fixed time, one might hope to get a more accurate estimate of position by designing a filter which minimizes the r. m. s. position uncertainty without regard for velocity uncertainty. It will be shown that the estimator which minimizes the r. m. s. uncertainty in the state vector also minimizes the r. m. s. uncertainty of any sub-vector of the state vector, and thus no improvement in the estimate can be achieved by neglecting the variation in velocity.

This concept can be extended to show that the same estimate is obtained even if the six integration constants being considered consist of the components of position variation at one time (e. g., the time of arrival at the destination) and the components of velocity variation at a different time (e. g., the "current" time).

3. Methods of Designing Estimators

The conventional methods of designing estimators for a multi-dimensional random variable fall into two main categories - the method of maximum likelihood and the technique based on optimum filter theory.

The maximum likelihood method was developed by statisticians and is based on the concept of maximizing a particular conditional probability. Practical estimators can be obtained from the method only by assuming that the individual measurements are normally distributed.

The object of optimum filter theory is to find a linear estimator which minimizes some function of the variances and covariances of the uncertainties in the estimated state vector. The development of the theory does not depend on the assumption that the uncertainties in the measurements are normally distributed.

Irrespective of whether maximum likelihood theory or optimum filter theory is used, it is possible to obtain either an "unbiased" or a "biased" estimate. An unbiased estimate is one in which only the direct measurements are used and no assumption is made about the a priori distribution of the uncertainties in the components of the state vector. The unbiased estimator is characterized by the fact that the estimated state vector coincides with the true state vector if all measurements are perfect. A biased, or Bayes, estimate is one in which the a priori distribution of the uncertainties in the components of the state vector is taken into account.

4. Optimization Criteria

There are three general types of criteria that have been used in applying optimization techniques to the redundant data problem. The first type, which is relevant to the method of maximum likelihood, consists of determining that state vector \underline{x} which maximizes the probability density of the occurrence of the measurements that have actually been obtained. This probability density is known as the likelihood function.

The second type of criterion involves minimizing the volume of the error ellipsoid associated with some linear function of the uncertainties in the components of the estimated state vector. If \underline{e} is

the n-dimensional column vector containing these uncertainties and if the vector \underline{w} is given by

$$\underline{w} = \overset{*}{R} \underline{e} \quad (4-1)$$

then the error ellipsoid in \underline{w} -space is defined by the equation

$$\underline{w}^T \overset{*}{W}^{-1} \underline{w} = 1 \quad (4-2)$$

where $\overset{*}{W}$ is the covariance matrix associated with \underline{w} .

$$\overset{*}{W} = \langle \underline{w} \quad \underline{w}^T \rangle \quad (4-3)$$

For example, if \underline{w} is equal to \underline{e} , $\overset{*}{R}$ becomes the identity matrix; if \underline{w} consists of only that portion of \underline{e} containing the uncertainties in position, $\overset{*}{R}$ is the following 3-by-6 matrix.

$$\overset{*}{R} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (4-4)$$

The third type of criterion minimizes the mean value of a positive semidefinite quadratic form in the uncertainties of the estimated state vector. In mathematical notation, the quantity to be minimized is

$$b = \langle \underline{e}^T \overset{*}{Q} \underline{e} \rangle \quad (4-5)$$

where $\overset{*}{Q}$ is a positive semidefinite matrix.

5. Mathematical Model

The number of measurements being processed is designated k; these measurements are combined in a k-dimensional column vector \underline{m} . In order to justify the use of linear theory, the components of \underline{m} are the deviations of the measured values from those values which would be observed ideally if the vehicle were on its reference trajectory.

In the general case \underline{x} , the parameter vector to be estimated, is of dimension n. Since it is assumed that redundant data are available, k is greater than n. For the state vector in space navigation n is equal to six.

The vector \underline{m} is assumed to be linearly related to \underline{x} in a manner that is a known function of the reference trajectory and of the type and time of each measurement.

$$\underline{m} = \underline{H}^* \underline{x} \quad (5-1)$$

\underline{H}^* is a k-by-n matrix whose elements are pre-computed. The rank of \underline{H}^* must be equal to n, for, if this were not so, \underline{x} could not be determined from \underline{m} .

The observed value of the measurement vector is $\underline{\tilde{m}}$, which differs from the true value due to the uncertainties in the individual measurements. The difference between $\underline{\tilde{m}}$ and \underline{m} is equal to the measurement uncertainty vector \underline{u} .

$$\underline{\tilde{m}} = \underline{m} + \underline{u} = \underline{H}^* \underline{x} + \underline{u} \quad (5-2)$$

The covariance matrix of the measurement uncertainties is the symmetric k-by-k matrix \underline{U}^* .

$$\underline{U}^* = \langle \underline{u} \underline{u}^T \rangle \quad (5-3)$$

For any practical case \underline{U}^* is a positive definite matrix.

The estimate of the state vector obtained from any of the filters to be developed is designated $\hat{\underline{x}}$. The error in the estimate is \underline{e} .

$$\underline{e} = \hat{\underline{x}} - \underline{x} \quad (5-4)$$

The covariance matrix of the components of \underline{e} is \underline{E}^* .

$$\underline{E}^* = \langle \underline{e} \underline{e}^T \rangle \quad (5-5)$$

6. Maximum Likelihood Theory

The likelihood function $L(\underline{x})$, in mathematical terms, is

$$L(\underline{x}) = p(\underline{\tilde{m}} | \underline{x}) \quad (6-1)$$

where $p(\underline{\tilde{m}} | \underline{x})$ is the probability density of the observed measurement vector $\underline{\tilde{m}}$, given the state vector \underline{x} . From Equation (5-2), $\underline{\tilde{m}}$ is the sum of a deterministic function of \underline{x} and the random vector \underline{u} . Therefore,

$$p(\underline{\tilde{m}} | \underline{x}) = p(\underline{u} | \underline{x}) \quad (6-2)$$

The probability density of \underline{u} is independent of \underline{x} , so that

$$L(\underline{x}) = p(\underline{u}) \quad (6-3)$$

$p(\underline{u})$ is the joint probability density of the components of \underline{u} . The components of \underline{u} are assumed to have a joint Gaussian distribution. Then,

$$L(\underline{x}) = \frac{\exp(-\frac{1}{2} \underline{u}^T \underline{U}^{*-1} \underline{u})}{[(2\pi)^k |\underline{U}^*|]^{1/2}} \quad (6-4)$$

where $|\underline{U}^*|$ is the determinant of \underline{U}^* .

Since $\log L(\underline{x})$ is a monotonically increasing function of $L(\underline{x})$, maximizing the logarithm with respect to \underline{x} yields the same value of \underline{x} as maximizing the likelihood function itself.

$$\begin{aligned} \log L(\underline{x}) &= -\frac{1}{2} \log \left[(2\pi)^k |\underline{U}^*| \right] - \frac{1}{2} \underline{u}^T \underline{U}^{*-1} \underline{u} \\ &= -\frac{1}{2} \log \left[(2\pi)^k |\underline{U}^*| \right] - \frac{1}{2} (\underline{\tilde{m}} - \underline{H}^* \underline{x})^T \underline{U}^{*-1} (\underline{\tilde{m}} - \underline{H}^* \underline{x}) \end{aligned} \quad (6-5)$$

The partial derivative of $\log L(\underline{x})$ with respect to x_i , one of the components of \underline{x} , when all other components of \underline{x} are held constant, is

$$\begin{aligned} \frac{\partial \log L(\underline{x})}{\partial x_i} &= \frac{1}{2} \left[\frac{\partial \underline{x}^T}{\partial x_i} \underline{H}^{*T} \underline{U}^{*-1} (\underline{m} - \underline{H}^* \underline{x}) \right. \\ &\quad \left. + (\underline{m} - \underline{H}^* \underline{x})^T \underline{U}^{*-1} \underline{H}^* \frac{\partial \underline{x}}{\partial x_i} \right] \end{aligned} \quad (6-6)$$

The second term on the right-hand side of Equation (6-6) is the transpose of the first term. Since both terms are scalar quantities, they are equal to each other.

$$\frac{\partial \log L(\underline{x})}{\partial x_i} = \frac{\partial \underline{x}^T}{\partial x_i} \underline{H}^{*T} \underline{U}^{*-1} (\underline{\tilde{m}} - \underline{H}^* \underline{x}) \quad (6-7)$$

The partial derivative $\frac{\partial \underline{x}^T}{\partial x_i}$ is a row vector with a one in the i-th position and all other components zero. Therefore, the right-hand side of Equation (6-7) represents the i-th component of the column vector $\underline{H}^{*T} \underline{U}^{*-1} (\underline{\tilde{m}} - \underline{H}^* \underline{x})$. When the likelihood function is a maximum, each of the n partial derivatives of $\log L(\underline{x})$ with respect to the components of \underline{x} vanishes. Then,

$$\underline{H}^{*T} \underline{U}^{*-1} (\underline{\tilde{m}} - \underline{H}^* \underline{x}) = \underline{0} \quad (6-8)$$

The solution of (6-8) for \underline{x} is $\hat{\underline{x}}_{ML}$, the maximum likelihood estimate of \underline{x} .

$$\hat{\underline{x}}_{ML} = \underline{F}_{ML}^* \underline{\tilde{m}} \quad (6-9)$$

where

$$\underline{F}_{ML}^* = (\underline{H}^{*T} \underline{U}^{*-1} \underline{H}^*)^{-1} \underline{H}^{*T} \underline{U}^{*-1} \quad (6-10)$$

It may be noted that

$$\overset{*}{F}_{ML} \overset{*}{H} = \overset{*}{I} \quad (6-11)$$

so that the maximum likelihood filter $\overset{*}{F}_{ML}$ is an unbiased filter.

From Equations (5-2) and (6-9),

$$\begin{aligned} \hat{\underline{x}}_{ML} &= \overset{*}{F}_{ML} \overset{*}{H} \underline{x} + \overset{*}{F}_{ML} \underline{u} \\ &= \underline{x} + \overset{*}{F}_{ML} \underline{u} \end{aligned} \quad (6-12)$$

The error \underline{e}_{ML} in the maximum likelihood estimate is

$$\underline{e}_{ML} = \hat{\underline{x}}_{ML} - \underline{x} = \overset{*}{F}_{ML} \underline{u} \quad (6-13)$$

The covariance matrix $\overset{*}{E}_{ML}$ is then given by

$$\begin{aligned} \overset{*}{E}_{ML} &= \langle \underline{e}_{ML} \underline{e}_{ML}^T \rangle \\ &= \overset{*}{F}_{ML} \overset{*}{U} \overset{*}{F}_{ML}^T \\ &= (\overset{*}{H}^T \overset{*}{U}^{-1} \overset{*}{H})^{-1} \end{aligned} \quad (6-14)$$

7. Gauss-Markov Theorem

If $\overset{*}{A}$ and $\overset{*}{B}$ are symmetric matrices, $\overset{*}{A}$ will be said to be greater than $\overset{*}{B}$ if the matrix difference $(\overset{*}{A} - \overset{*}{B})$ is positive definite. Similarly $\overset{*}{A} \geq \overset{*}{B}$ will indicate that $(\overset{*}{A} - \overset{*}{B})$ is positive semidefinite. The standard definitions of "positive definite" and "positive semidefinite" are given in Appendix A.

It is important to note that $\overset{*}{A} > \overset{*}{B}$ does not imply that each element of $\overset{*}{A}$ is greater than the corresponding element of $\overset{*}{B}$. For example,

$$\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} > \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

even though the off-diagonal elements of the right-hand matrix are larger than the off-diagonal elements of the left-hand matrix.

The Gauss-Markov theorem states that under the assumptions postulated in Section 5 $\overset{*}{E}_{ML}$ is less than or equal to the $\overset{*}{E}$ matrix for any other unbiased estimator. This theorem indicates that the variance of the component of the estimation uncertainty in any given direction in state space is minimized by the maximum likelihood estimate.

To prove the theorem, let $\overset{*}{F}_A$ be an arbitrary unbiased filter. The corresponding state

vector $\hat{\underline{x}}_A$ is

$$\hat{\underline{x}}_A = \overset{*}{F}_A (\overset{*}{H} \underline{x} + \underline{u}) \quad (7-1)$$

Since the filter is unbiased,

$$\overset{*}{F}_A \overset{*}{H} = \overset{*}{I} \quad (7-2)$$

Then, the estimation error \underline{e}_A is

$$\underline{e}_A = \hat{\underline{x}}_A - \underline{x} = \overset{*}{F}_A \underline{u} \quad (7-3)$$

The covariance matrix $\overset{*}{E}_A$ is given by

$$\overset{*}{E}_A = \langle \underline{e}_A \underline{e}_A^T \rangle = \overset{*}{F}_A \overset{*}{U} \overset{*}{F}_A^T \quad (7-4)$$

If $\overset{*}{J}$ is defined by

$$\overset{*}{J} = \overset{*}{F}_A - \overset{*}{F}_{ML} \quad (7-5)$$

the product $\overset{*}{J} \overset{*}{H}$ is

$$\overset{*}{J} \overset{*}{H} = \overset{*}{F}_A \overset{*}{H} - \overset{*}{F}_{ML} \overset{*}{H} = \overset{*}{I} - \overset{*}{I} = \overset{*}{O} \quad (7-6)$$

Substituting $\overset{*}{F}_A = \overset{*}{J} + \overset{*}{F}_{ML}$ into (7-4) yields

$$\begin{aligned} \overset{*}{E}_A &= \overset{*}{J} \overset{*}{U} \overset{*}{J}^T + \overset{*}{J} \overset{*}{U} \overset{*}{F}_{ML}^T \\ &\quad + (\overset{*}{J} \overset{*}{U} \overset{*}{F}_{ML}^T)^T + \overset{*}{F}_{ML} \overset{*}{U} \overset{*}{F}_{ML}^T \end{aligned} \quad (7-7)$$

From Equations (6-10) and (7-6),

$$\overset{*}{J} \overset{*}{U} \overset{*}{F}_{ML}^T = \overset{*}{J} \overset{*}{H} (\overset{*}{H}^T \overset{*}{U}^{-1} \overset{*}{H})^{-1} = \overset{*}{O} \quad (7-8)$$

From Equations (6-14), (7-7), and (7-8),

$$\overset{*}{E}_A - \overset{*}{E}_{ML} = \overset{*}{J} \overset{*}{U} \overset{*}{J}^T \quad (7-9)$$

In Appendix B it is shown that, since $\overset{*}{U}$ is positive definite, $\overset{*}{J} \overset{*}{U} \overset{*}{J}^T$ is positive semidefinite. Therefore,

$$\overset{*}{E}_A \geq \overset{*}{E}_{ML} \quad (7-10)$$

Because $\overset{*}{U}$ is positive definite, $\overset{*}{J} \overset{*}{U} \overset{*}{J}^T$ can be zero only if $\overset{*}{J}$ is the zero matrix. Consequently, $\overset{*}{E}_A$ can be equal to $\overset{*}{E}_{ML}$ only if $\overset{*}{F}_A$ is equal to $\overset{*}{F}_{ML}$.

It may be noted that in this derivation it has not been necessary to assume that the measurement uncertainties have a joint Gaussian distribution.

8. Minimizing the Volume of the Error Ellipsoid

The first type of optimum filter to be developed is that which minimizes the volume of the

error ellipsoid in \underline{w} -space. The equation of the error ellipsoid is (4-2). Its volume is

$$V = \frac{\pi^{\frac{q}{2}} |\underline{W}^*|^{1/2}}{\Gamma\left(\frac{q}{2} + 1\right)} \quad (8-1)$$

where \underline{w} is a q -dimensional vector. It is apparent that, for a given value of q , the volume of the ellipsoid is a function of the square root of the determinant of the covariance matrix \underline{W}^* . Thus, minimizing the ellipsoidal volume is equivalent to minimizing $|\underline{W}^*|$.

By the use of Equations (4-1) and (4-3) \underline{W}^* may be written in terms of \underline{R}^* and \underline{E}^* .

$$\underline{W}^* = \langle \underline{w} \underline{w}^T \rangle = \underline{R}^* \underline{E}^* \underline{R}^{*T} \quad (8-2)$$

If \underline{E}_{ML}^* and \underline{E}_A^* are, respectively, the covariance matrices associated with the uncertainties in the maximum likelihood estimator and with any other unbiased estimator, then from the Gauss-Markov theorem ($\underline{E}_A^* - \underline{E}_{ML}^*$) is positive semidefinite. In Appendix B it is shown that, if ($\underline{E}_A^* - \underline{E}_{ML}^*$) is positive semidefinite, the matrix product $\underline{R}^* (\underline{E}_A^* - \underline{E}_{ML}^*) \underline{R}^{*T}$, which is equal to ($\underline{W}_A^* - \underline{W}_{ML}^*$), is also positive semidefinite. Finally, from Appendix C,

$$|\underline{W}_A^*| \geq |\underline{W}_{ML}^*| \quad (8-3)$$

Therefore, use of the maximum likelihood estimator results in the minimum volume of the error ellipsoid in \underline{w} -space.

It may be shown that if $\underline{W}_A^* \geq \underline{W}_{ML}^*$ the error ellipsoid corresponding to the maximum likelihood estimator is contained in the error ellipsoid corresponding to the other estimator.

If the estimator \underline{F}_A^* actually minimizes $|\underline{W}^*|$ so that

$$|\underline{W}_A^*| = |\underline{W}_{ML}^*| \quad (8-4)$$

then from Appendix C the two \underline{W}^* matrices are equal.

$$\underline{W}_A^* = \underline{W}_{ML}^* \quad (8-5)$$

Thus,

$$\underline{R}^* \underline{E}_A^* \underline{R}^{*T} = \underline{R}^* \underline{E}_{ML}^* \underline{R}^{*T} \quad (8-6)$$

If $q = n$, so that \underline{R}^* is of rank n and can be inverted,

$$\underline{E}_A^* = \underline{E}_{ML}^* \quad (8-7)$$

In Section 7 it was shown that Equation (8-7) implies that

$$\underline{F}_A^* = \underline{F}_{ML}^* \quad (8-8)$$

Therefore, when $q = n$, the optimum filter is unique.

If $q < n$, there may be a family of filters \underline{F}_A^* , only one of which is equal to \underline{F}_{ML}^* , all satisfying the criterion that \underline{W}^* be minimized.

One possible type of \underline{R}^* matrix for the case $q = n$ is the state transition matrix \underline{C}_{ji}^* which transforms the state vector \underline{x}_i at time t_i to the state vector \underline{x}_j at time t_j .

$$\underline{x}_j = \underline{C}_{ji}^* \underline{x}_i \quad (8-9)$$

Then \underline{e} is the error vector associated with \underline{x}_i , and \underline{w} correspondingly is the error vector associated with \underline{x}_j . Since $q = n$, the unique filter which minimizes the volume of the six-dimensional error ellipsoid associated with the state vector \underline{x}_i is also the unique filter which minimizes the volume of the six-dimensional error ellipsoid associated with any other state vector \underline{x}_j . Thus the time to which the state vector is referred does not affect the accuracy of estimation.

If the \underline{R}^* matrix is a hybrid state transition matrix which transforms \underline{x}_i into a vector consisting of the position variation at one time and the velocity variation at another time or into position variations at two different times or into velocity variations at two different times, the same estimate of the state vector \underline{x}_s at some arbitrary time t_s is obtained as long as \underline{R}^* is nonsingular.

If the error ellipsoid to be minimized is associated with only the position components of some state vector, $q < n$, and in that case the maximum likelihood filter is only one of a set of possible optimum filters.

9. Minimizing Quadratic Forms in the Estimation Uncertainty

In Section 4 it was pointed that one of the criteria used in applying optimization techniques is minimizing the statistic

$$b = \langle \underline{e}^T \underline{Q} \underline{e} \rangle \quad (9-1)$$

Several versions of the positive semidefinite matrix $\overset{*}{\mathbf{Q}}$ are of practical importance. If $\overset{*}{\mathbf{Q}}$ is the identity matrix, the square root of b is the r. m. s. value of the magnitude of the error vector \underline{e}_i associated with the state vector \underline{x}_i . If $\overset{*}{\mathbf{Q}}$ is equal to $\overset{*}{\mathbf{C}}_{ji}^T \overset{*}{\mathbf{C}}_{ji}$, the square root of b is the r. m. s. value of the magnitude of the error vector \underline{e}_j associated with \underline{x}_j . If $\overset{*}{\mathbf{Q}}$ is given by

$$\overset{*}{\mathbf{Q}} = \begin{bmatrix} \overset{*}{\mathbf{I}} & \overset{*}{\mathbf{O}} \\ \overset{*}{\mathbf{O}} & h \overset{*}{\mathbf{I}} \end{bmatrix} \quad (9-2)$$

where h is a non-negative constant, the minimization involves velocity variation components that are weighted differently from the position variation components. In particular, if $h = 0$, the velocity components are ignored. It is also possible to choose $\overset{*}{\mathbf{Q}}$ such that b is equal to the mean of the sum of the squares of the position estimation uncertainty at one time and the velocity estimation uncertainty at another time.

Since $\underline{e}^T \overset{*}{\mathbf{Q}} \underline{e}$ is a scalar (a 1-by-1 matrix), it is equal to its trace. Then (9-1) becomes

$$b = \langle \text{tr} (\underline{e}^T \overset{*}{\mathbf{Q}} \underline{e}) \rangle \quad (9-3)$$

In Reference 7 it is shown that if $\overset{*}{\mathbf{A}}$ is an r -by- s matrix and $\overset{*}{\mathbf{B}}$ is an s -by- r matrix,

$$\text{tr} (\overset{*}{\mathbf{B}} \overset{*}{\mathbf{A}}) = \text{tr} (\overset{*}{\mathbf{A}} \overset{*}{\mathbf{B}}) \quad (9-4)$$

Then,

$$\begin{aligned} b &= \langle \text{tr} [(\underline{e}^T \overset{*}{\mathbf{Q}} \underline{e})] \rangle = \langle \text{tr} (\underline{e} \underline{e}^T \overset{*}{\mathbf{Q}}) \rangle \\ &= \text{tr} [\langle \underline{e} \underline{e}^T \rangle \overset{*}{\mathbf{Q}}] = \text{tr} (\overset{*}{\mathbf{E}} \overset{*}{\mathbf{Q}}) \end{aligned} \quad (9-5)$$

It will now be shown that the maximum likelihood filter $\overset{*}{\mathbf{F}}_{ML}$ minimizes the statistic b .

The Gauss-Markov theorem shows that $(\overset{*}{\mathbf{E}}_A - \overset{*}{\mathbf{E}}_{ML})$ is positive semidefinite. In Appendix D it is shown that if $(\overset{*}{\mathbf{E}}_A - \overset{*}{\mathbf{E}}_{ML})$ and $\overset{*}{\mathbf{Q}}$ are both positive semidefinite,

$$\text{tr} [(\overset{*}{\mathbf{E}}_A - \overset{*}{\mathbf{E}}_{ML}) \overset{*}{\mathbf{Q}}] \geq 0 \quad (9-6)$$

Thus,

$$\text{tr} (\overset{*}{\mathbf{E}}_A \overset{*}{\mathbf{Q}}) \geq \text{tr} (\overset{*}{\mathbf{E}}_{ML} \overset{*}{\mathbf{Q}}) \quad (9-7)$$

or

$$b_A \geq b_{ML} \quad (9-8)$$

When filter $\overset{*}{\mathbf{F}}_A$ minimizes b , so that b_A is equal to b_{ML} , then

$$\text{tr} [(\overset{*}{\mathbf{E}}_A - \overset{*}{\mathbf{E}}_{ML}) \overset{*}{\mathbf{Q}}] = 0 \quad (9-9)$$

If in addition $\overset{*}{\mathbf{Q}}$ is positive definite, then from Appendix D,

$$\overset{*}{\mathbf{E}}_A = \overset{*}{\mathbf{E}}_{ML} \quad (9-10)$$

From Section 7, Equation (9-10) implies that

$$\overset{*}{\mathbf{F}}_A = \overset{*}{\mathbf{F}}_{ML} \quad (9-11)$$

This equation indicates that the optimum filter is unique as long as $\overset{*}{\mathbf{Q}}$ is positive definite.

In Sections 6, 8, and 9, three different optimization criteria have been investigated. It has been shown that, regardless of which of the three criteria is used, the same optimum filter is obtained. This statement is true when any linearly independent set of six integration constants is used to define the vehicle's variant motion.

10. Biased and Unbiased Estimates

In this section it is proved that an unbiased estimator based on n fictitious measurements in addition to the k actual measurements can be constructed with the same covariance matrix $\overset{*}{\mathbf{E}}$ as a given biased estimator based on the k actual measurements. Conversely, a biased estimator can be constructed which has the same covariance matrix as a given unbiased estimator in which the n fictitious measurements have been taken into account. As a consequence of these facts, there is an optimum biased estimator $\overset{*}{\mathbf{F}}_{OB}$ which simultaneously minimizes the volumes of all error ellipsoids and also the means of all positive semidefinite quadratic forms in \underline{e} .

Biased estimators are important because they utilize available information in addition to that which is obtained from the direct measurements. The additional information is in the form of the first and second moments of the a priori uncertainty of the state vector \underline{x} .

When the optimum biased estimator is used, the resulting estimate of \underline{x} is more accurate than that which can be obtained from the optimum unbiased estimator $\overset{*}{\mathbf{F}}_{ML}$, in the sense that $\overset{*}{\mathbf{E}}_{OB} \leq \overset{*}{\mathbf{E}}_{ML}$ and $b_{OB} \leq b_{ML}$.

The measurement deviation vector which relates only to the measurements actually made will

be designated \underline{m}_B for biased estimation. The k-by-n matrix relating \underline{m}_B to \underline{x} is $\overset{*}{H}_B$.

$$\underline{m}_B = \overset{*}{H}_B \underline{x} \quad (10-1)$$

Correspondingly, the measurement uncertainty vector is \underline{u}_B , and its covariance matrix is $\overset{*}{U}_B$. $\overset{*}{F}_B$ is the biased filter that is to be used.

$$\hat{\underline{x}}_B = \overset{*}{F}_B \underline{\tilde{m}}_B \quad (10-2)$$

where $\hat{\underline{x}}_B$ is the biased estimate of the state vector. It is assumed a priori that the mean value of \underline{x} is the zero vector. The error vector for $\hat{\underline{x}}_B$ is \underline{e}_B , which is defined by

$$\underline{e}_B = \hat{\underline{x}}_B - \underline{x} \quad (10-3)$$

$\overset{*}{E}_B$ is the covariance matrix associated with \underline{e}_B .

The unbiased measurement deviation vector which includes the fictitious measurements is \underline{m}_U .

$$\underline{m}_U = \begin{bmatrix} \underline{m}_B \\ \underline{x} \end{bmatrix} \quad (10-4)$$

The fictitious measurements are assumed to be the components of the state vector \underline{x} . The observed unbiased measurement vector is

$$\underline{\tilde{m}}_U = \begin{bmatrix} \underline{\tilde{m}}_B \\ \underline{0} \end{bmatrix} \quad (10-5)$$

The (k+n)-by-n matrix $\overset{*}{H}_U$ relates \underline{m}_U to \underline{x} .

$$\underline{m}_U = \overset{*}{H}_U \underline{x} \quad (10-6)$$

where

$$\overset{*}{H}_U = \begin{bmatrix} \overset{*}{H}_B \\ \overset{*}{I} \end{bmatrix} \quad (10-7)$$

The measurement uncertainty vector \underline{u}_U is

$$\begin{aligned} \underline{u}_U &= \underline{\tilde{m}}_U - \underline{m}_U \\ &= \begin{bmatrix} \underline{u}_B \\ -\underline{x} \end{bmatrix} \end{aligned} \quad (10-8)$$

The covariance matrix $\overset{*}{U}_U$ is

$$\begin{aligned} \overset{*}{U}_U &= \langle \underline{u}_U \underline{u}_U^T \rangle \\ &= \begin{bmatrix} \overset{*}{U}_B & -\langle \underline{u}_B \underline{x}^T \rangle \\ -\langle \underline{x} \underline{u}_B^T \rangle & \langle \underline{x} \underline{x}^T \rangle \end{bmatrix} \end{aligned} \quad (10-9)$$

The unbiased estimate $\hat{\underline{x}}_U$ and the unbiased filter $\overset{*}{F}_U$ satisfy the equation

$$\hat{\underline{x}}_U = \overset{*}{F}_U \underline{\tilde{m}}_U \quad (10-10)$$

$\overset{*}{F}_U$ is an n-by-(k+n) matrix, while $\overset{*}{F}_B$ is an n-by-k matrix. If $\overset{*}{F}_B$ is given, then

$$\overset{*}{F}_U = \begin{bmatrix} \overset{*}{F}_B & \overset{*}{F}_R \end{bmatrix} \quad (10-11)$$

where the residual matrix $\overset{*}{F}_R$ is an n-by-n matrix. It will now be shown that if $\overset{*}{F}_U$ is to be unbiased, then

$$\overset{*}{F}_R = \overset{*}{I} - \overset{*}{F}_B \overset{*}{H}_B \quad (10-12)$$

For an unbiased estimate,

$$\overset{*}{F}_U \overset{*}{H}_U = \overset{*}{I} \quad (10-13)$$

From Equations (10-7) and (10-11),

$$\begin{aligned} \overset{*}{F}_U \overset{*}{H}_U &= \begin{bmatrix} \overset{*}{F}_B & \overset{*}{F}_R \end{bmatrix} \begin{bmatrix} \overset{*}{H}_B \\ \overset{*}{I} \end{bmatrix} \\ &= \overset{*}{F}_B \overset{*}{H}_B + \overset{*}{F}_R \end{aligned} \quad (10-14)$$

When (10-13) and (10-14) are solved for $\overset{*}{F}_R$, Equation (10-12) is obtained.

By Equation (7-4), the unbiased covariance matrix $\overset{*}{E}_U$ is given by

$$\overset{*}{E}_U = \overset{*}{F}_U \overset{*}{U}_U \overset{*}{F}_U^T \quad (10-15)$$

The uncertainty vector \underline{e}_B can be written in terms of the unbiased filter characteristics by utilizing Equations (10-8), (10-11), and (10-12).

$$\underline{e}_B = \overset{*}{F}_B \overset{*}{H}_B \underline{x} + \overset{*}{F}_B \underline{u}_B - \underline{x}$$

$$= \overset{*}{F}_B \underline{u}_B - \overset{*}{F}_R \underline{x}$$

$$= \overset{*}{F}_U \begin{bmatrix} \underline{u}_B \\ -\underline{x} \end{bmatrix} = \overset{*}{F}_U \underline{u}_U \quad (10-16)$$

Finally,

$$\overset{*}{E}_B = \langle \underline{e}_B \quad \underline{e}_B^T \rangle = \overset{*}{F}_U \overset{*}{U}_U \overset{*}{F}_U^T \quad (10-17)$$

By comparing (10-15) with (10-17), it is seen that

$$\overset{*}{E}_U = \overset{*}{E}_B \quad (10-18)$$

By reversing the argument presented in Equations (10-11) through (10-18), it may be shown that if $\overset{*}{F}_U$ is given and $\overset{*}{F}_B$ is obtained by deleting the last n columns of $\overset{*}{F}_U$, then the biased and unbiased covariance matrices obtained satisfy Equation (10-18).

Thus, irrespective of whether one starts with a biased or unbiased filter, a filter of the other type can be constructed such that the two filters have identical $\overset{*}{E}$ matrices.

Deleting the last n columns of the maximum likelihood filter $\overset{*}{F}_{ML}$ obtained with the use of fictitious measurements yields the optimum biased filter $\overset{*}{F}_{OB}$. Then,

$$\overset{*}{E}_{OB} \leq \overset{*}{E}_B \quad (10-19)$$

where $\overset{*}{E}_B$ is the covariance matrix associated with the estimation uncertainty obtained by the use of any other n -by- k filter $\overset{*}{F}_B$. This is true because there is an n -by- $(n+k)$ unbiased filter $\overset{*}{F}_B$ with covariance matrix E_B , and by the Gauss-Markov theorem

$$E_B \geq E_{ML} = E_{OB} \quad (10-20)$$

Therefore, all the expressions developed in Sections 8 and 9 relating optimum unbiased filters to other unbiased filters are also applicable to the relationship between optimum biased filters and other biased filters.

APPENDIX A

A symmetric r -by- r matrix $\overset{*}{M}$ is said to be positive definite if the quadratic form

$$\underline{x}^T \overset{*}{M} \underline{x} > 0 \quad (A-1)$$

for every non-zero vector \underline{x} . If $\underline{x}^T \overset{*}{M} \underline{x}$ is greater than or equal to zero but not necessarily greater than zero for every non-zero vector \underline{x} , $\overset{*}{M}$ is said to be positive semidefinite.

The principal subdeterminants of a matrix may be used to test whether or not the matrix is positive definite. The principal subdeterminant m_k of $\overset{*}{M}$ is defined as the determinant of the k -by- k matrix formed from the elements in the first k rows and the first k columns of $\overset{*}{M}$. For example, if the element in the i -th row and j -th column of $\overset{*}{M}$ is designated M_{ij} ,

$$m_1 = M_{11} \quad (A-2)$$

$$m_2 = \begin{vmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{vmatrix} \quad (A-3)$$

and

$$m_r = \left| \overset{*}{M} \right| \quad (A-4)$$

$\overset{*}{M}$ is positive definite if and only if all of its principal subdeterminants are positive. If some of the principal subdeterminants are zero but none are negative, $\overset{*}{M}$ is positive semidefinite.

APPENDIX B

Theorem: If $\overset{*}{A}$ is a positive semidefinite r -by- r matrix and $\overset{*}{B}$ is an arbitrary r -by- s matrix, then the s -by- s matrix $\overset{*}{B}^T \overset{*}{A} \overset{*}{B}$ is positive semidefinite.

Proof: Let \underline{z} be an arbitrary s -dimensional vector and let

$$\underline{y} = \overset{*}{B} \underline{z} \quad (B-1)$$

Then

$$\underline{y}^T \overset{*}{A} \underline{y} = \underline{z}^T \overset{*}{B}^T \overset{*}{A} \overset{*}{B} \underline{z} \geq 0 \quad (B-2)$$

because $\overset{*}{A}$ is positive semidefinite. Hence $\overset{*}{B}^T \overset{*}{A} \overset{*}{B}$ is positive semidefinite.

APPENDIX C

Theorem: Let $\overset{*}{A}$ and $\overset{*}{B}$ be symmetric r-by-r matrices such that

$$\overset{*}{A} \geq \overset{*}{B} > \overset{*}{O} \quad (C-1)$$

Then,

$$|\overset{*}{A}| \geq |\overset{*}{B}| > 0 \quad (C-2)$$

Also,

$$|\overset{*}{B}| = |\overset{*}{A}| \quad (C-3)$$

only if

$$\overset{*}{B} = \overset{*}{A} \quad (C-4)$$

Proof: By a matrix theorem used in the theory of vibrations (Page 47 of Reference (8)), $\overset{*}{A}$ and $\overset{*}{B}$ can be simultaneously diagonalized by a non-singular matrix $\overset{*}{M}$.

$$\overset{*}{M}^T \overset{*}{A} \overset{*}{M} = \overset{*}{D} \quad (C-5)$$

$$\overset{*}{M}^T \overset{*}{B} \overset{*}{M} = \overset{*}{I} \quad (C-6)$$

where $\overset{*}{D}$ is a diagonal matrix whose i-th diagonal element is d_i . If $\overset{*}{N}$ is the inverse of $\overset{*}{M}$,

$$\overset{*}{A} - \overset{*}{B} = \overset{*}{N}^T (\overset{*}{D} - \overset{*}{I}) \overset{*}{N} \quad (C-7)$$

Let \underline{h}_k be a column vector whose k-th component is one and whose other components are all zeros. Let

$$\underline{p}_k = \overset{*}{M} \underline{h}_k \quad (C-8)$$

Then

$$\underline{p}_k^T (\overset{*}{A} - \overset{*}{B}) \underline{p}_k = (d_k - 1) \geq 0 \quad (C-9)$$

The inequality holds because $(\overset{*}{A} - \overset{*}{B})$ is positive semidefinite. From (C-9),

$$d_k \geq 1 \quad (C-10)$$

for $k = 1, \dots, r$.

The determinants of $\overset{*}{B}$ and $\overset{*}{A}$ are given by

$$|\overset{*}{B}| = |\overset{*}{N}|^2 \quad (C-11)$$

$$|\overset{*}{A}| = d_1 d_2 \dots d_r |\overset{*}{N}|^2 = d_1 d_2 \dots d_r |\overset{*}{B}| \quad (C-12)$$

In consequence of (C-10), (C-11), and (C-12),

$$|\overset{*}{A}| \geq |\overset{*}{B}| > 0 \quad (C-13)$$

If $|\overset{*}{B}| = |\overset{*}{A}|$, $d_1 = d_2 = \dots = d_r = 1$, and from (C-7)

$$\overset{*}{A} - \overset{*}{B} = \overset{*}{O} \quad (C-14)$$

APPENDIX D

Theorem: Let $\overset{*}{A}$ and $\overset{*}{B}$ be positive semi-definite symmetric r-by-r matrices, then

$$\text{tr}(\overset{*}{A} \overset{*}{B}) \geq 0 \quad (D-1)$$

If $\overset{*}{A}$ is positive semidefinite and $\overset{*}{B}$ is positive definite, then

$$\text{tr}(\overset{*}{A} \overset{*}{B}) = 0 \quad (D-2)$$

only if

$$\overset{*}{A} = \overset{*}{O} \quad (D-3)$$

Proof: Since $\overset{*}{A}$ is positive semidefinite, it can be diagonalized by an orthogonal matrix $\overset{*}{M}$.

$$\overset{*}{D} = \overset{*}{M}^T \overset{*}{A} \overset{*}{M} \quad (D-4)$$

where $\overset{*}{D}$ is a diagonal matrix whose i-th diagonal element is d_i . The fact that $\overset{*}{A}$ is positive semidefinite implies that

$$d_i \geq 0 \quad (D-5)$$

for $i = 1, \dots, r$. Let $\overset{*}{D}_1$ be a diagonal matrix whose i-th diagonal element is $d_i^{1/2}$. Then

$$\overset{*}{D} = \overset{*}{D}_1 \overset{*}{D}_1^T \quad (D-6)$$

Now, let

$$\overset{*}{A}_1 = \overset{*}{M} \overset{*}{D}_1 \overset{*}{M}^T \quad (D-7)$$

Because $\overset{*}{M}$ is orthogonal, it follows that

$$\begin{aligned} \overset{*}{A}_1 \overset{*}{A}_1^T &= \overset{*}{M} \overset{*}{D}_1 \overset{*}{M}^T \overset{*}{M} \overset{*}{D}_1^T \overset{*}{M}^T \\ &= \overset{*}{M} \overset{*}{D} \overset{*}{M}^T = \overset{*}{A} \end{aligned} \quad (D-8)$$

Since $\text{tr}(\overset{*}{A} \overset{*}{B}) = \text{tr}(\overset{*}{B} \overset{*}{A})$,

$$\begin{aligned} \text{tr}(\overset{*}{A} \overset{*}{B}) &= \text{tr}[\overset{*}{A}_1 (\overset{*}{A}_1^T \overset{*}{B})] \\ &= \text{tr}(\overset{*}{A}_1^T \overset{*}{B} \overset{*}{A}_1) \end{aligned} \quad (D-9)$$

By Appendix B, the matrix $\overset{*}{A}_1^T \overset{*}{B} \overset{*}{A}_1$ is positive semidefinite and therefore has non-negative diagonal elements. Thus, in view of (D-9),

$$\text{tr}(\overset{*}{A} \overset{*}{B}) \geq 0 \quad (D-10)$$

To prove the second part of the theorem, let $\overset{*}{A}_1$ be partitioned into column vectors $\underline{s}_1, \dots, \underline{s}_r$.

$$\overset{*}{A}_1 = [\underline{s}_1 \dots \underline{s}_r] \quad (D-11)$$

Then if $\text{tr}(\overset{*}{A} \overset{*}{B}) = 0$, it follows that

$$0 = \text{tr}(\overset{*}{A}_1^T \overset{*}{B} \overset{*}{A}_1) = \text{tr}(\overset{*}{A}_1^T [\overset{*}{B} \underline{s}_1 \dots \overset{*}{B} \underline{s}_r])$$

$$= \sum_{i=1}^r \underline{s}_i^T \overset{*}{B} \underline{s}_i \quad (\text{D-12})$$

Since $\overset{*}{B}$ is positive definite,

$$\underline{s}_i^T \overset{*}{B} \underline{s}_i \geq 0 \quad (\text{D-13})$$

and hence

$$\underline{s}_i^T \overset{*}{B} \underline{s}_i = 0 \quad (\text{D-14})$$

by Equation (D-21). Again using the fact that $\overset{*}{B}$ is positive definite, it follows that

$$\underline{s}_i = \underline{0} \quad (\text{D-15})$$

for $i = 1, \dots, r_1$ thus

$$\overset{*}{A}_1 = \overset{*}{O} \quad (\text{D-16})$$

and by (D-8)

$$\overset{*}{A} = \overset{*}{A}_1 \overset{*}{A}_1^T = \overset{*}{O} \quad (\text{D-17})$$

References

1. McLean, J.D., Schmidt, S.F., and McGee, L.A., "Optimal Filtering and Linear Prediction Applied to a Midcourse Navigation System for the Circumlunar Mission," NASA Tech. Note D-1208, 1962.
2. Carr, R.E. and Hudson, R.H., "Tracking and Orbit Determination Program of the Jet Propulsion Laboratory," JPL Technical Report No. 32-7, Feb. 1960.
3. Smith, G.L., Schmidt, S.F., and McGee, L.A., "Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on board a Circumlunar Vehicle," NASA Tech. Rept. R-135, 1962.
4. Battin, R.H., "A Statistical Optimizing Navigation Procedure for Space Flight," ARS Journal, Vol. 32 No. 11, Nov. 1962, pp. 1681 - 1696.
5. Stern, R.G., "Interplanetary Midcourse Guidance Analysis," Sc. D. thesis in Department of Aeronautics and Astronautics, M.I.T., May 1963.

6. Noton, A.R.M., "Interplanetary Post-Injection Guidance," JPL External Publication No. 653, June 1959.
7. Bellman, R., "Introduction to Matrix Analysis," McGraw-Hill Book Co., New York, 1960.
8. Hildebrand, F.B., Methods of Applied Mathematics, Prentice-Hall, Inc., New York, 1952.